L1 norm constrained migration of blended data with the FISTA algorithm

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 J. Geophys. Eng. 12 620
(http://iopscience.iop.org/1742-2140/12/4/620)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 109.171.137.222
This content was downloaded on 14/09/2015 at 05:12

Please note that terms and conditions apply.
1. Introduction

Silverman (1979) proposed the concept of simultaneous shooting for vibroseis-type and impulsive sources and this notion was extended to incoherent shooting and blending (blended acquisition) by Berkhout (2008). Blended acquisition overcomes the disadvantages of traditional acquisition such as long acquisition cycle, low collection efficiency, and high cost. Blended acquisition is a great innovation to the traditional acquisition and promotes the progress of seismic acquisition technology.

In blended seismic acquisition, the time intervals between shots are small and different sources shoot in an overlapping fashion to acquire blended shot records. This shooting and recording strategy makes up for the shortcoming of the poor sampling but results in two or more shot records blending together and brings great difficulties to the follow-up migration processing. At present, there are two options with regard to the migration of blended data. In option 1, a preprocessing procedure—called ‘deblending’, which separates the blended data into a single shot data, is needed. Then standard migration processing flows are applied to these deblended data. In option 2, migration processing is directly applied to the blended data without any pre-separation. In order to improve the processing efficiency and reduce the impact of deblending, we use the direct migration methods. However, the direct imaging of blended data is not satisfactory due to the crosstalk contamination. Assuming that the distribution of subsurface reflectivity is sparse, in this paper, we formulate the seismic imaging problem of blended data as a Basis Pursuit denoise (BPDN) problem. Based on compressed sensing, we propose a L1 norm constrained migration method applying to the direct imaging of blended data. The Fast Iterative Shrinkage-Thresholding Algorithm, which is stable and computationally efficient, is implemented in our method. Numerical tests on the theoretical models show that the crosstalk introduced by blended sources is effectively attenuated and the migration quality has been improved enormously.

Keywords: blended acquisition, direct imaging, crosstalk noise, L1 norm, compressed sensing, fast iterative shrinkage-thresholding algorithm

(Some figures may appear in colour only in the online journal)
regularization method (Tikhonov and Arsenin 1977) implements a quadratic penalty item to weaken the original ill-posed problem. Although the quadratic regularization term can give the stable approximate solution, it tends to provide a solution that is smooth and dispersed, whereas the reflectivity distribution of the subsurface is sparse (Aldawood 2013).

Aldawood proposed the L1 norm constraint prestack Kirchhoff migration in order to preserve the sparsity of subsurface reflectivity and improve the imaging resolution. He utilized the SPG (Birgin et al 2000) (Spectral Projected Gradient method) to solve the large-scale sparse reconstruction problem. Numerical examples showed that reflectivity models obtained by using this L1 norm constraint algorithm were highly accurate with desired resolution.

Since the subsurface reflectivity distribution is sparse and discontinuous, based on the compressed sensing algorithm (Donoho 2006a, Baraniuk 2007), we formulate the seismic imaging problem of blended data which contains blending noise as a Basis Pursuit denoise (BPDN) problem (van den Berg and Friedlander 2008), extend the sparse constraint theory from the traditional acquisition system to the blended acquisition system, and propose a L1 norm constrained migration method applying to the direct migration of the blended data. Considering the cost and the quality of imaging, Fast Iterative Shrinkage-Thresholding Algorithm (Beck and Teboulle 2009) (FISTA) is implemented in our method. Numerical tests on the theoretical models show that the crosstalk introduced by blended sources is effectively attenuated and the migration quality has been significantly improved.

2. Theory

2.1. Blended acquisition and direct migration

Within the limits of the Born approximation, seismic data d can be generated by the forward Born modeling operator L and the reflectivity model m:

\[ d = Lm. \]  

(1)

For blended acquisition geometry, the phase-encoded blended data can be described as:

\[ \tilde{d} = \sum_{i=1}^{S} P_i d_i, \]  

(2)

where \( \tilde{d} \) is the blended data (i.e. supergather) consisting of \( S \) shots, \( P_i \) is the phase-encoding operator and \( d_i \) is the \( i \)th single shot gather. Substituting (1) into (2), we obtain:

\[ \tilde{d} = \sum_{i=1}^{S} P_i L_i m = \tilde{L} m, \]  

(3)

where

\[ \tilde{L} = \sum_{i=1}^{S} P_i L_i, \]  

(4)

is defined as the supergather modeling operator and \( L_i \) represents the forward modeling operator for \( d_i \). A composite diagram of one blended shot record in time domain is showing in figure 1.

The multisource blended datasets can be formulated as:

\[ [d_1 \ldots d_N][\tilde{P}_1 \ldots \tilde{P}_M] = [\tilde{d}_1 \ldots \tilde{d}_M]. \]  

(5)

The matrix equation (5) expresses that there are a total of \( N \) individual sources and every \( S = N / M \) sources are shot in an overlapping fashion with certain time delays to acquire \( M \) blended shot records. \( \tilde{P}_k \) is a \( N \times 1 \) column vector defining the array and the emission time delays of blended single sources in \( \tilde{d} \). Users can control the locations and time delays of blended sources by the design of \( \tilde{P} \) and in practice it is proposed that all the \( \tilde{P}_k \) are designed to make sure that \( \tilde{P}_k^T \tilde{P}_k \) is equal to the unity matrix.

The supergather migration operator is regarded as the adjoint of the supergather modeling operator:

\[ \tilde{L}^T = \sum_{i=1}^{S} L_i^T P_i^T. \]  

(6)

So the phase-encoded migration of supergather \( \tilde{m}_{mig} \) is obtained by applying the migration operator \( \tilde{L}^T \) to the blended data \( \tilde{d} \):

\[ \tilde{m}_{mig} = \tilde{L}^T \tilde{d}. \]  

(7)
From (3) and (6), we get:

$$\tilde{m}_{\text{mig}} = L^T d = L^T L m.$$  

(8)

This equation shows that the migration imaging is a $$L^2$$ blurred version of the actual reflectivity model $$m$$, which normally has an acquisition footprint and arc phenomenon. In order to improve the quality of the migration image, we can regard the migration imaging as a seismic inversion problem.

At present, discussing migration imaging in the framework of seismic inversion, the least-square algorithm plays the role as the bridge, that is, reconstructing the reflectivity model $$m$$ to fit the observed data $$d$$ in L2 norm:

$$J(m) = \frac{1}{2}||Lm - d||_F^2.$$  

(9)

However, the above equation is a typical ill-posed problem in mathematical physical inverse problems, the objective function of which needs a regularization constraint to restrict the morbid degree. At this stage, LSM normally uses L2 norm regularization. In this paper, we use the L1 norm regularization method for optimization and a stable and sparse solution can be obtained.

### 2.2. L1 norm constrained migration method

The sparse representation of migration can be represented as follows:

$$\min ||m||_0, \text{ subject to } d = Lm.$$  

(10)

Minimizing the L0 norm of $$m$$ leads to the sparsity of $$m$$. However, it has been proven that this L0 norm optimization problem is a typical NP hard problem (Candes et al. 2006), which is unstable and has a large numerical computation. Donoho (2006b) proved that under the specified conditions, both the L0 norm minimization and L1 norm minimization had the same unique solution. Furthermore, the L1 norm minimization is a convex optimization problem that can turn the NP hard problem into a linear programming problem. In practical processing, the common and tractable way is to use the L1 norm as a sparse constraint condition:

$$\min ||m||_1, \text{ subject to } d = Lm.$$  

(11)

This L1 norm constraint problem is also known as a Basis Pursuit (BP) problem (Chen et al. 1998). Because of the cross-talk noise in the migration of the blended data, we formulate the migration problem as a Basis Pursuit denoise (BPDN) problem as follows:

$$\min ||m||_1, \text{ subject to } ||Lm - d||_2 \leq \sigma (\sigma \text{ is a non-negative small scalar}),$$  

(12)

where $$\sigma$$ quantifies the noise level in the measurements. This formulation takes priority when a reasonable estimate of $$\sigma$$ can be known. We rephrase the BPDN problem in Lagrangian form, which has a close connection to the convex quadratic programming and can use various solution algorithms. So the penalty function of the imaging problem is defined as follows:

$$J(m) = \frac{1}{2}||Lm - d||_E^2 + \lambda ||m||_1,$$  

(13)

Equation (13) consists of two terms: the first term is the data misfit function used to measure the error between the observed and the fitting data and the second term is a L1 norm regularization term used to penalize non-sparse solutions. $$\lambda$$ is a non-negative regularization parameter, which determines the amount of sparsity desired in the model. It is important to select an appropriate value of $$\lambda$$. When $$\lambda$$ is large, the fitting data is sufficiently sparse, but the solution may not be properly fit the observed data. Conversely, when $$\lambda$$ is too small, the fitting data may honor the observed data but the solution would be severely overestimated.

In this paper, we use the Fast Iterative Shrinkage-Thresholding Algorithm to minimize (13). FISTA develops on the Iterative Shrinkage-Thresholding Algorithms (Figueiredo and Nowak 2003, Daubechies et al. 2004) (ISTA). It is attractive due to its simplicity and high computing efficiency, and thus is adequate for solving large-scale data.

The gradient iteration for solving (13) is of the form:

$$m_{k+1} = m_k + \frac{1}{\alpha}(d - Lm_k),$$  

(14)

where $$\alpha$$ is a suitable steps size controlling the iterative step, i.e. convergence speed.

The (14) can be equivalent to a linearized function at $$m_{k-1}$$ and written as:

$$m_k = \arg \min_m \left\{ \frac{\alpha}{2} ||m - m_{k-1} - \frac{1}{\alpha} \nabla J(m_{k-1})||_1^2 + \lambda ||m||_1 \right\}.$$  

(15)

The shrinkage or soft-threshold operator is widely employed in the image denoising domain and it is the core of ISTA/FISTA. For the general case, the shrinkage operator is defined as:

$$\text{soft}(x, T) = \begin{cases} x - T & x > T \\ 0 & -T \leq x \leq T, \\ x + T & x < -T \end{cases}$$  

(16)

where T is the threshold. The main function of the shrinkage operator is shifting the parameter x towards the origin with T units, i.e. returning to zero (figure 2).

Solving (15) in the class of ISTA leads to the iterative scheme:
\[ \mathbf{m}_{k+1} = \text{soft}\left( \mathbf{m}_k + \frac{1}{\alpha} \mathbf{L}^H (\mathbf{d} - \mathbf{L} \mathbf{m}_k), \frac{\lambda}{2\alpha} \right) \]  

where each iteration involves a matrix-vector multiplication involving \( \mathbf{L} \) and \( \mathbf{L}^H \) followed by a soft threshold step, and each iteration gradually narrows the difference between the previous iteration result and the back projection. Through multiple iterations, the \( \mathbf{m}_k \) would converge to a constant value. A typical condition ensuring the convergence of \( \mathbf{m}_k \) to a minimizer \( \mathbf{m} \) of (13) is to require that \( \alpha \geq \max \text{eig}(\mathbf{L}^H \mathbf{L}) \), i.e. the stepsize is greater than the maximum eigenvalue of \( \mathbf{L}^H \mathbf{L} \frac{\lambda}{2\alpha} \) is the threshold.

The development in FISTA is using a linear combination of the results of the two previous iterations as the beginning of current iteration. The new intermediate parameter \( \mathbf{y}_k \) is as follows:

\[ \mathbf{y}_{k+1} = \mathbf{m}_k + \left( \frac{t_k - 1}{t_{k+1}} \right) (\mathbf{m}_k - \mathbf{m}_{k-1}), \]  

where \( t_k \) is defined as:

\[ t_{k+1} = \frac{1 + \sqrt{1 + 4 t_k^2}}{2}. \]

Using \( \mathbf{y}_k \) instead of \( \mathbf{m}_k \) in (17), we obtain the general iterative step of FISTA:

\[ \mathbf{m}_{k+1} = \text{soft}\left( \mathbf{y}_k + \frac{1}{\alpha} \mathbf{L}^H (\mathbf{d} - \mathbf{L} \mathbf{y}_k), \frac{\lambda}{2\alpha} \right). \]

The basic realization flow of FISTA is as follows:

Initial value assignment: \( \mathbf{y}_1 = \mathbf{m}_0, t_1 = 1. \)

Step \( k \):

1. \( \mathbf{m}_{k+1} = \text{soft}\left( \mathbf{y}_k + \frac{1}{\alpha} \mathbf{L}^H (\mathbf{d} - \mathbf{L} \mathbf{y}_k), \frac{\lambda}{2\alpha} \right) \).
2. \( t_{k+1} = \frac{1 + \sqrt{1 + 4 t_k^2}}{2} \).
3. \( \mathbf{y}_{k+1} = \mathbf{m}_k + \left( \frac{t_k - 1}{t_{k+1}} \right) (\mathbf{m}_k - \mathbf{m}_{k-1}) \).

3. Numerical tests

3.1. Migration for a scatter model

First, we compute the standard migration, LSM and L1 norm constrained migration for a simple scatter model shown in figure 3. 9 scattering points having a reflectivity of 1 are evenly distributed in the homogeneous medium with a background velocity of 1500 m s\(^{-1}\). The scatter model has 48 $\times$ 48 grids in the \( x \) and \( z \) dimensions. 48 sources and 48 receivers are deployed on the surface. A Ricker wavelet with a dominant frequency of 30 Hz is used in this experiment.

To simulate a variety of multisource blended datasets, the 48 conventional individual shots are encoded and blended together to form 3 different clusters of blended datasets: 12 supergathers with 4 individual shots per supergather (figure 4(a)), 6 supergathers with 8 individual shots per supergather (figure 4(b)) and 2 supergathers with 24 individual shots per supergather (figure 4(c)). The blending code of each dataset is a simple random short time delay between 0 and 1 s. All the random time delays are generated by a random number generator in Matlab. We use the Kirchhoff modeling and migration operator for our tests, which has the advantages of a fast calculating speed and strong adaptability for the seismic observation system. Our experiments use the Matlab parallel programming on a computer with Intel(R) Core(TM) i5 processor running at 3.40 G Hz and 16 G RAM. The imaging results of the scatter model can be seen in figure 5.

Figures 5(a), (d) and (g) show the Kirchhoff migration images which are heavily blurred by notable migration artifacts and bad focus. Also, these images indicate that the crosstalk noise is becoming more serious with the decrease in the number of supergathers. Figures 5(b), (e) and (h) show the LSM images and figures 5(c), (f) and (i) are the L1 norm constrained migration images. Compared with the standard migration, the LSM reduces some of the crosstalk and migration artifacts. However, the imaging results of the LSM still have notable noise and inaccurate focus. The sparse solutions of the L1 norm constrained migration show that our method can completely remove the migration artifacts and crosstalk and have better-focused results of the reflected energy. The calculation parameters used in FISTA and computing time of L1 norm constrained migration and LSM are listed in table 1.

As can be seen from table 1, when the model is very sparse, the L1 norm constrained migration using FISTA has less computing time to get better imaging results than the LSM does.

3.2. Migration for the 2D Marmousi model

In order to verify our method is also effective for a complex model, it is tested on the 2D Marmousi model. To reduce the computation burden, we rarely the original model without lowering the computational accuracy and the model complexity. After re-sampling, the model has 300 $\times$ 150 grids in the \( x \) and \( z \) dimensions with the grid interval of 10 m. 150 sources and 300 receivers are deployed on the surface. Other test parameters of the seismic observation system are the same as the
Figure 4. One of the blended shot gathers with (a) 4 single sources, (b) 8 single sources and (c) 24 single sources.

Figure 5. (a) Kirchhoff migration for 12 4-shot supergathers, (b) LSM for 12 4-shot supergathers, (c) L1 norm constrained migration for 12 4-shot supergathers, (d) Kirchhoff migration for 6 8-shot supergathers, (e) LSM for 6 8-shot supergathers, (f) L1 norm constrained migration for 6 8-shot supergathers, (g) Kirchhoff migration for 2 24-shot supergathers, (h) LSM for 2 24-shot supergathers and (i) L1 norm constrained migration for 2 24-shot supergathers.
previous scatter model. Figure 6 shows the velocity model and reflectivity model calculated from the velocity model using vertical rays and constant density assumptions.

In this section, the 150 conventional individual shots are blended together to form 3 clusters of blended datasets: 30 supergather with 5 individual shots per supergather (figure 7(a)), 10 supergather with 15 individual shots per supergather (figure 7(b)), and 3 supergather with 50 individual shots per supergather (figure 7(c)). The migration imaging results obtained from the 3 blended datasets and the waveform comparison can be seen in figure 8–10.

Figures 8(a), 9(a) and 10(a) show the Kirchhoff migration images of the blended data. It is obvious that the crosstalk and migration artifacts worsen the image quality. With the number of supergathers decreasing, the Kirchhoff migration image of the blended data becomes terrible and the seismic events are almost submerged by noise. Figures 8(b), 9(b) and 10(b) show the L1 norm constrained migration images of the blended data. Compared with the conventional Kirchhoff migration images, the L1 norm constrained migration images become clear and the noise is efficiently attenuated. Moreover, complicated geologic structures such as anticline, fault, and thin interbed get the desired imaging results and the fuzzy events become continuous and well-distributed. Figures 8(c), 9(c) and 10(c) plot the waveform comparison of the blended data migration images and the true reflection coefficients. The comparison data is the 250th trace. The waveform comparison further proves the validity of the proposed method in the micro-scale. Through these comparisons we can see that the L1 norm constrained migration results approximate well to the true reflection coefficients and the crests and troughs have significant consistency. On the other hand, the standard migration image

<table>
<thead>
<tr>
<th>Blended datasets</th>
<th>L1 norm constrained migration</th>
<th>LSM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda$</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>12 4-shot supergathers</td>
<td>0.05</td>
<td>0.5</td>
</tr>
<tr>
<td>6 8-shot supergathers</td>
<td>0.05</td>
<td>0.5</td>
</tr>
<tr>
<td>2 24-shot supergathers</td>
<td>0.05</td>
<td>10</td>
</tr>
</tbody>
</table>

Figure 6. (a) 2D Marmousi model (velocity). (b) 2D Marmousi model (reflectivity).

Figure 7. One of the blended shot gathers with (a) 5 single sources, (b) 15 single sources, and (c) 50 single sources.
Figure 8. Imaging results for 30 5-shot supergathers from (a) Kirchhoff migration, (b) L1 norm constrained migration and (c) the waveform comparison of true reflectivity (red line), L1 norm constrained migration image trace (green line) and Kirchhoff migration (blue line) image trace (the data is the 250th trace).

Figure 9. Imaging results for 10 15-shot supergathers from (a) Kirchhoff migration, (b) L1 norm constrained migration and (c) the waveform comparison of true reflectivity (red line), L1 norm constrained migration image trace (green line) and Kirchhoff migration (blue line) image trace (the data is the 250th trace).
trace is very different from the other two traces. The calculation parameters and computing time of the L1 norm constrained migration in the 2D Marmousi model experiments are listed in Table 2.

4. Conclusion

In this paper, we formulate the migration problem of the blended data as a BPDN and propose a L1 norm constrained migration method applying to direct migration of the blended data. Considering the computational cost and the quality of imaging, the FISTA is utilized for the optimization.

The tests of the simple model and the complex Marmousi model both show that the method put forward in this paper has a better imaging results for the blended data. In the final imaging results, the crosstalk and migration artifacts are effectively attenuated, the migration images are accurately converged to the true location of the reflectors, and the migration quality is significantly improved. The L1 norm constrained migration method is applicable to the common migration algorithms such as wave equation migration and reverse time migration, and the computational efficiency and accuracy of the results depend on the selection of the migration operators.

Finally, there are a few points to note:

As for the model type, when the subsurface reflection coefficient model is sufficient sparse, the superiority of the L1 norm constrained migration algorithm is obvious. It can produce high-quality imaging results with lower computational cost.

During the migration iterations, the improper choice of source wavelet can impair the quality of migration images. In this article, we assume that the estimation of the seismic wavelet is correct and use the same wavelet in forward and migration. The L1 norm regularization term can also eliminate some of the effects of the source wavelet. In real seismic data processing, the source wavelet can be estimated by stacking the near-offset ocean-bottom reflections.

Acknowledgments

We would like to thank the China National Nature Science Foundation (No. 41374115) and the National ‘863’ Project (No. 2014AA06A605) for their funding support.
References

Aldawood A, Hoteit I and Alkhalifah T 2013 Compressed-sensing application-pre-stack Kirchhoff migration EAGE Conf. and Exhibition Incorporating SPE EUROPEC (London)


Berkhout A J 2008 Changing the mindset in seismic data acquisition The Leading Edge 27 924–38


Dai W, Fowler P and Schuster G T 2012 Multi-source least-squares reverse time migration Geophys. Prospect. 60 681–95

Daubechies I, Defrise M and Mol C D 2004 An iterative thresholding algorithm for linear inverse problems with a

sparsity constraint Commun. Pure Appl. Math. 57 1413–541


Donoho D L 2006b For most large underdetermined systems of linear equations the minimal L1-norm solution is also the sparest solution Commun. Pure Appl. Math. 59 797–829


Nemeth T, Wu C and Schuster G T 1999 Least-squares migration of incomplete reflection data Geophysics 64 208–21


Silverman D 1979 Method of three dimensional seismic prospecting US Patent 4, 159, 463


Tikhonov A N and Arsenin V Y 1977 Solutions of Ill-Posed Problems (Washington, DC: V. H. Winston and Sons) pp 20–40